Identifying User Engagement Patterns in an Online Video Discussion Platform

Seung Yeon Lee
EdLab, Teachers College
Columbia University
tmddus30@gmail.com

Hui Soo Chae
EdLab, Teachers College
Columbia University
hsc2001@tc.columbia.edu

Gary Natriello
EdLab, Teachers College
Columbia University
gjn6@columbia.edu

ABSTRACT

In this study we conducted behavioral analyses to gain insights into patterns of user interaction in a video discussion platform, *Vialogues*. Vialogues provides an asynchronous online discussion environment around video. Using a hierarchical clustering analysis on users' clickstream data, we identified four different behavior patterns: (1) video watchers with no discussion activity, (2) opinion seekers and active repliers with little to no video watching activity, (3) users who watched and discussed videos, and (4) users focused on viewing and/or creating metadata. Despite being the largest group, Cluster (3) had the least classifiable characteristics. Consequently we conducted additional analyses to examine finer-grained user segments. For each segment we created a transition network using weighted directed networks in order to understand the transition pattern between two consecutive click activities.

Keywords

Online discussion, video learning, hierarchical clustering, transition network, user behavior

1. INTRODUCTION

Using video as an instructional technology has been a popular approach across a broad range of educational contexts. Video provides a way for learners to immediately connect with subject matter; increasingly, it also provides a way for learners to connect with each other. A number of benefits to using video in education have been reported over several decades of research [7, 8, 10, 11, 5]. While traditional video platforms primarily support passive learning, social video platforms, (e.g., YouTube), provide an active learning environment for learners to discuss video and share content collaboratively.

Vialogues [1] is an asynchronous online video discussion platform that facilitates collaborative conversations around video. The platform allows users to comment directly on specific points in time of a video, as opposed to commenting only in the comment section that references the entire video. The main video is shown on the left side of the screen, and the discussion board is shown on the right side of the screen, as shown in Figure 1. As described above, all comments are coded to a specific point in time in the video, and the related portions of the video are referenced. Also, the discussions are threaded so that users are able to view and respond to one another's comments. The addition of this feature in Vialogues allows deeper understanding of video by enabling users to understand the context and to discuss via conversation threads; this resolves one of the main problems of many existing video-discussion tools.

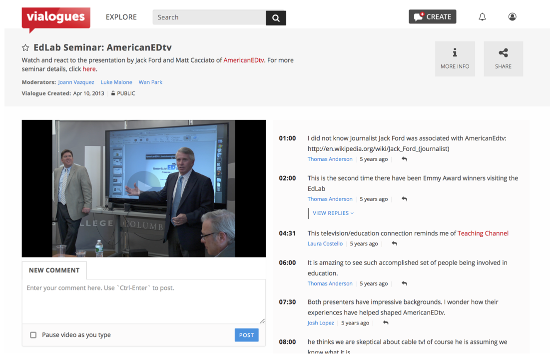

Figure 1: An example vialogue page. The page shows the title and description of the vialogue (top), the video player (left) and discussion panel (right).

Vialogues provides a comprehensive set of pedagogical tools to assist teachers to flexibly design and monitor learning activities, as well as receive feedback from students based on instructional needs. Teachers can ask either survey questions, with a 'check all that apply' answer option or devise poll questions, with a 'single answer' option. Teachers are able to present these questions throughout different points in the video. In addition, users can either open their discussion to everyone, or they can restrict access to some specific users. With these additional tools, discussion moderators can effectively control the quality of their discussion and tailor a discussion for an intended group of users.

In this paper, we conducted behavioral analyses to gain insights into users' interaction patterns in Vialogues. We performed several types of analyses on users' clickstream data in a step wise way to develop deeper understanding of user

Table 1: Four event categories on Vialogues, (1) video player, (2) video watch, (3) discussion, (4) other features, and corresponding event actions of each event category

Video Player	Video Watch	Discussion			Other Features
		Comments	Reply	Poll	
Video ready	Watch 3 seconds	Post comments	Click reply	Post poll	Pause as typing
Play	Watch 10 seconds	Click edit comment	Reply	Add poll item	Click time code
Pause	Watch 30 seconds	Cancel edit comment	Expand reply	Remove poll item	Open Vialogues tab
Mute true	Watch 50%	Update comment	Hide reply		Close Vialogues tab
Mute false	Watch 95%	Click delete comment			Open settings tab
Fullscreen true	Watch 100%	Delete comment			Save edit Vialogue
Fullscreen false					Save edit Video
					Cancel edit Vialogue

engagement. First, we examined the overall usage patterns based on the distribution of different user actions. Then we investigated various interaction patterns by using a hierarchical clustering analysis. The clustering analysis identified four user groups and we conducted in-depth analysis on each groups to understand its distinctive behavior characteristics. We found that different groups demonstrated different levels of engagement, particularly in terms of discussion activity. The findings of this study could be used to create a useful reference for designing instruction based on video discussion tools or for developing this kind of learning platforms.

2. DATA

2.1 Data Source

We analyzed the clickstream actions generated as users interact with Vialogues. We considered four event categories to understand users' behaviors on a particular vialogue page (e.g., Figure 1). First, we collected users' actions interacting with the video player, e.g., play, pause, mute, full screen. Second, we recorded the video completion rate, e.g, watch the first 3 seconds, watched 50% of the total duration of video, etc. Third, we collected actions related to discussion. For context, Vialogues supports three different ways to participate in a discussion: commenting, replying and posting polls. Actions related to each of these three activities were collected, e.g, post comments, reply to others' comments, expand replies, post a poll, etc. Lastly, other actions were also tracked including whether they used a "Pause as Typing" feature which automatically pauses the video while typing comments; whether they clicked the "Time Code" in the discussion panel to find the corresponding part of video; and whether they opened the vialogues tab to see more information about the particular vialogue such as the uploader or the shareable link. Table 1 presents the full list of tracked actions under each of the four categories described above, which we defined as 'Video Player', 'Video Watch', 'Discussion (Comments, Reply, Poll)' and 'Other Features' respectively. We used the data of users who visited a vialogue page during the month of September 2017.

2.2 Data Pre-processing

We conducted data cleaning and exploratory analyses to preprocess the original sample data. First, we excluded vialogue contents created by Vialogues administrators. Second, we examined the number of different vialogue pages a user visited within a single session in order to understand its distribution and detect any outliers; this numbers varied from 1 to 37 but 95% of the data had values between 1 and 3. We only considered those 95% of the data and assumed that the data with values greater than 3 were outliers. We believe that when a user explores too many contents within the same session, he or she is less likely to be fully engaged in watching videos or participating in discussions. Next, we processed data to create a vector for the sequence of actions for each case, where cases were defined as a unique vialogue page within a session for each user. For example, if a user interacted with two different vialogue pages during a given session, two vectors were defined for this user. We treated them as separate cases because users' behaviors can not only be different by sessions but also by vialogue contents. From this process, we created 6,706 distinct cases from the September data. In other words, the total unique sets of sessions and vialogue pages viewed in September 2017 were 6,706. In vialogue pages, by the system setting, users' first action were logged as "Video Ready". However, 563 cases out of the total 6,706 started with different event actions other than video ready. This indicates that some users could interact with the same vialogue pages through multiple sessions. For example, users may pause their interaction with Vialogues and come back to the same page after some period of time. The event actions of those aforementioned sessions are dependent with the actions of the corresponding previous sessions. Thus, for the purpose of our analysis, these 563 cases were excluded as they could not be treated as independent cases like the others. We also examined the length of action sequences. We found that its distribution is extremely skewed: 90% of the cases were between 1 and 50 and there were outliers with extreme values up to 357. After deleting these outliers, we also excluded the first action, video ready, from every vector. This event action was created by default, not by users' intention, and is present in every cases, so does not provide any useful information about the user engagement. After such data processing steps, the analytic sample from September 2017 included 3,485 unique cases of 2,972 sessions from 1,516 unique user IDs and 991 unique vialogue contents.

3. OVERALL PATTERN OF USER INTERACTION

In order to understand the overall pattern of users' behavior on vialogue pages, we first evaluated the frequencies of event actions in September 2017. This analytic sample, constructed by the process described above, included 33 event actions from the actions given in Table 1, and excludes "Video Ready" as described above. Table 2 shows the frequency count of the top 10 most frequent actions. The distribution suggests that the actions of playing or pausing

Table 2: Frequency of the top 10 most frequent actions

Event Action	Freq
Video pause	9,818
Video play	8,445
Expand reply	4,262
Video watch 3 Secs	3,070
Video watch 10 Secs	2,839
Video watch 30 Secs	2,575
Post comment	2,157
Video watch 50%	1,878
Video watch 95%	1,181
Click reply comment	1,179
Video fullscreen true	1,109

videos occurred the most frequently, which was expected. The next most frequent action, interestingly, was expanding replies. In the platform, users can view others' first-level comments without any click activity but they need to click to read replies to a particular comment. Although the system cannot capture the action of simply reading others' comments without clickable actions, "Expand reply" can serve as a proxy for the behavior of browsing others' discussion comments. Then, the events "Watch for 3 seconds," "Watch for 10 seconds," and "Watch for 30 seconds" were the next most frequent actions in that order. The next most frequent action was "Post comment" followed by "Video watch 50%" then "Video watch 95%." This shows that users are not necessarily watching the complete video before commenting. It also suggests that some users are commenting while watching a video. One interesting finding regarding replying to comments is that the count of "Reply comment" (freq = 907) is less than the count of "Click reply comment" (freq = 1,179). It is possible that some users tried to reply to a comment but ultimately decided not to post a reply. It is also possible that some users were confused by the platform's layout and clicked the button by mistake. Overall, it appears that viewers typically browse discussion contents, watch a video for a certain duration, and post comments (not replies) before completing a video. It also shows that people reply to comments (freq = 907) less frequently than completing 95% of a video (freq = 1,181). Note that these observations are at the aggregate level and do not take into consideration different user sessions or time sequences. However, our goal is not generalization. Instead our goal is to develop a better sense of overall user activity patterns.

4. IDENTIFYING DISTINCT BEHAVIOR PATTERNS

4.1 Method

To further investigate user groups with distinct behavior patterns, we conducted a clustering analysis. For clustering, we created a $M \times N$ matrix, where M is the total number of cases and N is the number of different actions. Columns indicate unique actions and rows indicate cases defined as unique vialogue pages visited within an individual session. The elements of each row are the frequency counts of each action in each case. In the data processing step, we found large variation in the total number of actions among different cases, and this may cause invalid clustering results. To minimize such risks, we normalized each row and then computed the distances between the pair of rows using the cosine similarity. The cosine similarity measures the similarity based on the angle between two vectors ignoring the frequency of each element. For two vectors, $\boldsymbol{a} = \{a_i\}$ and $\boldsymbol{b} = \{b_i\}$ for $i = 1, ..., M$, the cosine similarity is calculated by $Similarity = cos(\theta) = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{||\boldsymbol{a}||_2 ||\boldsymbol{b}||_2} = \frac{\sum_{i=1}^{M} a_i b_i}{\sqrt{\sum_{i=1}^{M} a_i^2}\sqrt{\sum_{i=1}^{M} b_i^2}}$. The corresponding distance was computed by $1 - Similarity$. Using the cosine-based distances, we applied Ward's hierarchical cluster method [6]. Specifically, we used the "agglomerative approach" which goes from the bottom up. This approach starts with each data point in a cluster of its own. Then, it repeats the process of finding the most similar pair of clusters and merging them until all data are merged into only one cluster. The Ward's method uses the minimum variance criterion which minimizes the total within-cluster variance: at each step, combine two clusters whose merge results in the smallest increase in the total within-cluster variation. We determined the optimal number of clusters based on the Calinski-Harabasz (CH) index [2]. The CH index is calculated by $\frac{SS_B/(k-1)}{SS_W/(N-k)}$, where k is the number of clusters and N is the total number of cases. SS_B is the total between-cluster variance, which measures how spread apart the groups are from each other; and SS_W is the total within-cluster variance, which measures high tightly grouped the clusters are. As the number of clusters increases, SS_B keeps increasing while SS_W keeps decreasing. The CH index finds the clustering assignment that simultaneously has a large SS_B and a small SS_W by using the variance ratio criterion; the largest CH index occurs with the optimal number of clusters. The analysis was conducted using `hclust()` in R.

In order to interpret each of the different clusters, we considered the proportion vector of actions for each case. This is calculated by the frequency of the particular action (e.g., play video) divided by the count of all actions. This allows for more intuitive interpretation than using the normalized vectors and still resolves the issue that arises from varying the total number of actions for different cases. For each of the resulting clusters, we then calculated the average of the above described proportions across all the cases that were assigned to the particular cluster. Based on these metrics, we interpreted each cluster to identify distinct patterns.

4.2 Results

The resulting dendrogram from the Ward's hierarchical clustering is shown in Figure 2. The computed CH index sug-

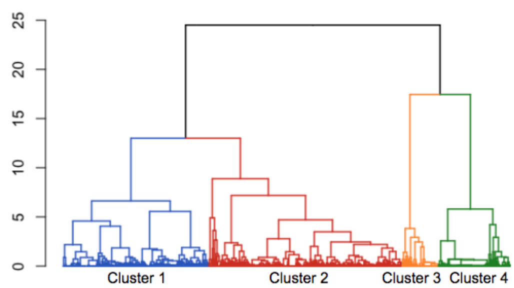

Figure 2: Dendrogram of results from the Ward's hierarchical cluster method

Figure 3: The four cluster profiles, or interaction patterns

gested four clusters as the best number of clusters. For each cluster, we examined the size of the cluster and the distribution of different actions based on the average proportion of each action per case, as described in a preceding method section. In order to understand the different patterns of each cluster, we plotted the average proportion of actions. Figure 3 illustrates each cluster's profile; in the x-axis, we listed different actions that exhibit similar characteristics. We first listed actions related to video watching behavior and labeled as "Watch (Video)": (in the listed order) video play, video pause, video mute true, video mute false, video full screen true, video full screen false, video watch 3 seconds, video watch 10 seconds, video watch 30 seconds, video watch 50%, video watch 95%, video watch 100%. Then, we listed those actions related to discussion activity and labeled it as "Read & Interact (Discussion)": (in the listed order) post comment, expand reply, hide reply, click time code, pause as typing, reply comment, click reply comment, click delete comment, click edit comment, cancel edit comment, update comment, delete comment, post poll, remove poll item, add poll item. As the last category, we listed actions related to viewing and creating/managing meta data and labeled as "View & Manage (Metadata)": (in the listed order) open vialogues tab, close vialogues tab, open settings tab, save edit vialogue, save edit video, cancel edit vialogue.

A graphical analysis of Figure 3 leads to the following observation: Cluster 1 shows high focus on video watching activities with no noticeable occurrence of other actions. In Cluster 2, the peaks, representing locally frequent actions, are somewhat spread out, but the graph shows the highest concentration in video watching and discussion activities. In Cluster 3, frequent actions are centered around the viewing and creating metadata. Cluster 4 shows a heavy focus on discussion activities with very limited number of other activities.

Based on the preliminary analysis of the graphs, we conducted additional examinations to understand each cluster. For the purpose of this examination, we assigned ranks based on the frequency of actions occurring in each cluster. Table 3 presents the top 10 most frequent actions on average for each cluster. In Cluster 1, the frequent actions were all related to the video watching activities: video play/pause, watch a video for a certain duration of time, and use of a full screen mode, which suggests that Cluster 1's dominating pattern is pure video watching. In Cluster 2, however, we observed that the discussion activities (post comments, expand reply) were also present in addition to video watching actions. These discussion activities were limited to the first-level interaction, and are more interactive than just watching video. However, there was limited interaction with other users/viewers. In other words, users in Cluster 2 were commenting on the video but not discussing the video with other users. Cluster 3 was unique in that the most frequent action was "Open vialogues tab" which is often used when users look for other information about the specific video such as the uploader, the upload date, and sharing features. The proportion of such actions was dominant at 0.64 while other actions' proportions were less than 0.1. Additionally, for this cluster, other actions related to editing and setting the vialogue contents were the next most frequent actions: save/edit vialogue, open settings tab, which are only allowed for content creators and moderators. Thus, the behaviors present in Cluster 3 predominantly consist of exploring the peripheral information and creating/editing vialogue metadata. In Cluster 4, the most frequent action was expanding others' replies, with the proportion of 0.69. This cluster showed the heightened focus on discussion activity in that eight of the top 10 actions were discussion-related: expand reply, click reply comment, hide reply, reply comment, click edit comment, update comment, cancel edit comment, post comment. The remaining two actions were video play/pause and no video watching for a certain period. Thus, Cluster 4 represents "opinion seeking" and "replying" behaviors.

In terms of the cluster size, the sizes were 1,137, 1,508, 282, 558 for Cluster 1, 2, 3, and 4, respectively. It is noteworthy that Cluster 2 is the largest cluster; 43% (= 1,508/3,485) of cases were assigned to Cluster 2, which was characterized as a mix of video watching and discussion activities. This behavior pattern, which combines both video and discussion, was expected to be the most popular pattern considering

Table 3: Top 10 most frequent actions and the averaged proportions in each cluster

Cluster 1 (size = 1,137)		Cluster 2 (size = 1,508)		Cluster 3 (size = 282)		Cluster 4 (size = 558)	
Video play	0.255	Video pause	0.292	Open vialogues tab	0.639	Expand reply	0.684
Video watch 3 Secs	0.144	Video play	0.187	Close vialogues tab	0.08	Click reply comment	0.071
Video watch 10 secs	0.129	Video watch 3 secs	0.08	Save edit vialogue	0.056	Hide reply	0.068
Video watch 30 secs	0.109	Video watch 10 secs	0.069	Open settings tab	0.051	Reply comment	0.044
Video pause	0.078	Post comment	0.067	Video play	0.039	Click edit comment	0.02
Video watch 50%	0.069	Video watch 30 secs	0.053	Video pause	0.028	Update comment	0.016
Video watch 95%	0.046	Expand reply	0.043	Video watch 3 secs	0.022	Cancel edit comment	0.014
Video watch 100%	0.04	Video watch 50%	0.033	Video watch 10 secs	0.017	Video play	0.012
Video full screen true	0.038	Video full screen true	0.02	Video watch 30 Secs	0.012	Video pause	0.011
Video full screen false	0.035	Video full screen false	0.018	Post comment	0.008	Post comment	0.01

that the key feature of Vialogues is its support of discussion around video content. Also noticeable was Cluster 3, which had the smallest size, with only 282 cases present (8% = 282/3,485). Cluster 3 consisted of exploring and creating metadata. Its small cluster size can be partially explained by the fact that most frequent actions of Cluster 3 were creating or editing activities available only to creators or moderators, not participants.

5. TRANSITION PATTERNS BETWEEN DIFFERENT EVENT ACTIONS

As the core objective of Vialogues is to promote discussion around video, it is important to evaluate the case in which users' usage patterns exhibit both video watching and discussion activities, e.g., identifying sequences of actions [4]. In the clustering analysis above, Cluster 2 was the largest group with both video watching and discussion, but did not show a clear classifiable pattern. Thus, we examined finer-grained user groups out of Cluster 2. In the Ward's clustering analysis, when the number of clusters increased to 6 (compared to 4 in the above analysis), Cluster 2 was further broken down into 3 clusters while Cluster 1, 3 and 4 remained as is. For the 3 sub-clusters generated from Cluster 2, using the same approach, the average proportions of actions were examined. In this case, however, we only examined the actions associated with video watching and discussion: video watch 3 secs, video watch 10 secs, video watch 30 secs, video watch 50%, video watch 95%, video watch 100%, click timecode, expand reply, post comment, and reply comment. Using only these 10 actions, the proportions of action frequency were recalculated for each action (i.e., the frequency of each action divided by the total number of frequency of 10 actions). Figure 4 presents profiles of each sub-clusters. Sub-cluster 1 represents modest amounts of both video watching and discussion, with a high proportion of posting comments among discussion activity. On the other hand, Sub-cluster 2 and 3 show heavier focus on discussion activities. Sub-cluster 2 had the highest proportion of "posting comment" action and Sub-cluster 3 had the highest peak at the "expanding replies" action.

We further examined each of the sub-clusters to understand the sequence of actions when users are involved in both video watching and discussing activities. We assumed that previous actions may potentially influence the following actions and sought to explore the path that a user takes to participate in discussion. We performed a transition network analysis, specifically applying weighted directed networks [12]. The benefit of this method is that we can discover the transition pattern between the two consecutive click activities and also gain insight about the degree by which the same action transition patterns appear. We generated weighted directed networks for each sub-cluster using the `ngram` [9] and `igraph` [3] packages in R.

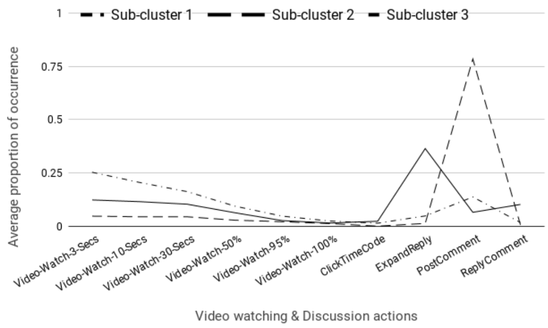

Figure 4: Profiles of three sub-clusters of Cluster 2

The networks for three sub-clusters are presented in Figure 5. Each action name was abbreviated as follows: 's' is the indication of start, 'V-S' indicates video watching for 3, 5, 10 seconds, 'V-50' includes video watching 50%, 95%, 100%, 'PC' indicates posting comments, 'CT' is clicking timecode, 'ER' indicates expanding others' replies, and 'RC' is replying to others comments, and 'e' indicates the end of the action. The directed edges indicate transition between two consecutive actions and the width (weight) of the edges indicates the number of that transition occuring. The node size for each action indicates the total number of frequency of the particular action in the aggregated sequence set.

The left network describes transition patterns of Sub-cluster 1. Since it was characterized by a combination of a modest amount of both video watching and discussion activities, the directed edges exist for various combinations of actions with similar weights. Specifically, the transition from 's' to 'V-S' had the largest weight, indicating many users in this sub-cluster started by watching a video rather than conducting other actions. Other frequent transitions were one from 'V-S' to 'V-50' and two transitions from each of the 'V-S' and 'V-50' to 'PC'. Overall, transitions between video watching and posting comments were dominant. The network for Sub-cluster 2 is presented in the middle. For this sub-cluster, the most frequent action was posting comments. Interestingly, the graph shows heavy weights on the edge from 's' to 'PC'

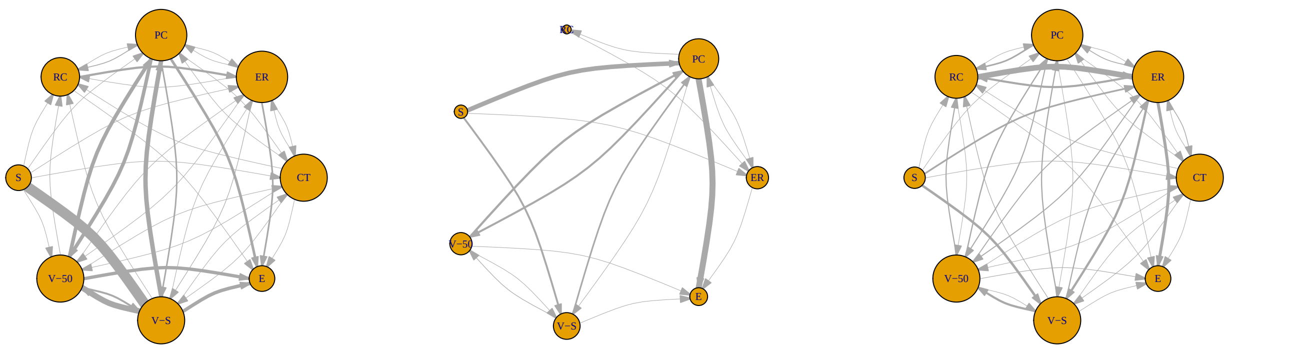

Figure 5: Weighted directed networks for Sub-cluster 1 (left), Sub-cluster 2 (middle) and Sub-cluster 3 (right)

and the one from 'PC' to 'e', which implies that users in this cluster tend to post their comment at the beginning even before watching a video and then leave the page. This could indicate that some users might read others' first level comments (which does not generate clickstream data in the current system) and then post their own comments. If this conjecture proves true, an interesting question for this group would be why are these users only posting first-level comments without replying to others' replies. This could be a future area of inquiry that helps to uncover users' path to interactive online discussion around video. Lastly, for Sub-cluster 3, the transitions from replying to comments (RC) to expanding other replies (ER) were noticeable. An interesting finding from this graph is that interactions with replies was not necessarily derived from video watching since we could not observe any significantly noticeable transition from video watching to interactive discussion with others. This suggests that for some group of users, others' comments or other factors that were not captured in clickstream data might have greater effects on replying behavior rather than the video itself.

6. CONCLUSION

In our study, we identified users' behavior patterns on Vialogues in an exploratory manner. It is important to note that while there exist a number of academic studies on the value of video based education, there are limited research papers that specifically deal with the discussion in the context of a video platform. This paper contributes to the field since it is focused on online video-based discussion, which was made possible through the Vialogues video discussion platform. Clustering analysis of different user group behaviors can provide a point of reference for future studies but more importantly, this can help educators to enhance video-based instruction and learning.

7. REFERENCES

[1] M. Agarwala, I. H. Hsiao, H. S. Chae, and G. Natriello. Vialogues: Videos and dialogues based social learning environment. In *2012 IEEE 12th International Conference on Advanced Learning Technologies*, pages 629–633, July 2012.

[2] T. Calinski and J. Harabasz. A dendrite method for cluster analysis. *Communications in Statistics – Theory and Methods*, 3(1):1–27, 1974.

[3] G. Csardi and T. Nepusz. The igraph software package for complex network research. *InterJournal*, Complex Systems:1695, 2006.

[4] D. Ferreira, M. Zacarias, M. Malheiros, and P. Ferreira. Approaching process mining with sequence clustering: Experiments and findings. In G. Alonso, P. Dadam, and M. Rosemann, editors, *Business Process Management*, pages 360–374, 2007.

[5] K. J. Lee and M. D. Sharma. Incorporating active learning with videos: A case study from physics. *Teaching Science: The Journal of the Australian Science Teachers Association*, 54(4), 2008.

[6] F. Murtagh and P. Legendre. Ward's hierarchical agglomerative clustering method: Which algorithms implement ward's criterion? *Journal of Classification*, 31(3):274–295, 2014.

[7] D. Passey. *Digital video technologies enhancing learning for pupils at risk and those who are hard to reach.*, pages 156–168. Glasgow Caledonian University Press, 2006.

[8] O. Poquet, L. Lim, N. Mirriahi, and S. Dawson. Video and learning: A systematic review (2007–2017). In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, LAK '18, pages 151–160, New York, NY, USA, 2018. ACM.

[9] D. Schmidt and C. Heckendorf. *Guide to the ngram Package: Fast n-gram Tokenization*, 2017. R Vignette.

[10] S. Schwan and R. Riempp. The cognitive benefits of interactive videos: learning to tie nautical knots. *Learning and Instruction*, 14(3):293–305, 2004.

[11] I. Vieira, A. P. Lopes, and F. Soares. The potential benefits of using videos in higher education. In *Proceedings of EDULEARN14 Conference*, pages 0750–0756. IATED Publications, 2014.

[12] S. Wasserman and K. Faust. *Social network analysis: Methods and applications*, volume 8. Cambridge university press, 1994.